

# **Building a Virtual Humanities Collaboratory**

**@ CRASSH, 17 Mill Lane, Cambridge**

**Tuesday, 6 January to Wednesday, 7 January**

CRASSH, 17 Mill Lane, Cambridge, CB2 1RX  
Tel: +44 1223 766886  
Fax: +44 1223 765276

## **The TEI and Semantic Interoperability**

Susan Schreibman

The Text Encoding Initiative is a standard, a research community, and an abstract model for describing textual material from ancient grave inscriptions to born digital text. This paper will survey the development of the TEI as a means for interchange, to an international standard for text encoding, to a framework for other methods to reuse and mine in the creation of new knowledge paradigms.

## **Modification, Reuse and Subversion: Digital Object Collections and the Humanities.**

Robin Boast

Collaboratories and e-Humanities, following the lead of Collaboratories and eScience, have assumed foundational models of stable, functional and tool-rich laboratories. However, 20 years of Sociology of Scientific Knowledge (Bruno Latour, John Law, Michelle Callon, Andrew Pickering, Harry Collins, etc.) have shown that the scientific laboratory is anything but stable, that its functionality is constantly renegotiated and its tools survive only to the degree that they are transformable. A major research programme at the MAA is constructively questioning how Digital Object Collections and Collaboration Spaces can be constructed to allow for modification, reuse and even subversion of digital resources.

## **Tibetan books and digital dilemmas: issues of standards and digital asset management in a collaborative project.**

Hildegard Diemberger

This paper will present a research project that has worked as a 'pilot' for the creation of a Union Catalogue of Tibetan resources by cataloging and partially digitizing Tibetan manuscripts and blockprints kept at the Cambridge University Library, the Bodleian, the British Library and other UK institutions. The catalogue is designed to provide an interactive device that will enable the progressive integration of other collections of Tibetan resources kept in the UK and relate to parallel projects in other part of the world. This endeavour has involved collaboration at different levels among people within the University of Cambridge, among different institutions in the UK and among international networks of expertise. The paper will outline this collaborative setting and address issues of standards and digital asset management encountered during the realisation of the project.

# **Collaboration, Access and other Unexpected Consequences: Reflections on 8 years of the Digital Himalaya Project**

Mark Turin

The Digital Himalaya Project <[www.digitalhimalaya.org](http://www.digitalhimalaya.org)> was established in December 2000 in the Department of Social Anthropology in Cambridge with three primary objectives:

1. to preserve in a digital medium anthropological materials from the Himalayan region that were fast degenerating in their current form, including film, photographs, audio recordings, field notes, maps and rare publications
2. to make these resources available over the WWW so that they might be used by researchers and students
3. to make these resources available on DVD to the descendants of the people from whom the materials were collected

Thanks to financial support secured in the United Kingdom and later in the United States, and through partnerships with larger digital library initiatives, we have succeeded in reaching our goals.

This presentation focuses on the unexpected turns that the Digital Himalaya Project has taken over the last 8 years, and concludes with a look to the future.

While the project began as a strategy for salvaging, archiving and disseminating the products of (primarily colonial) ethnographic collections on the Himalayas—both for posterity and for heritage communities—Digital Himalaya has now become a collaborative digital publishing environment which brings a new collection online every month. The website has grown from being a static homepage with occasional updates to a dynamic content delivery platform for over 20GB of archived data. Similarly, our website has moved from being almost exclusively used by members of Western universities to providing a range of services to a global public, with a particularly strong user base in Asia. Digitisation has been 'off-shored' to Nepal, dramatically reducing operational costs and increasingly productivity. And most of all, our modest funding no longer comes from national funding bodies in Europe or America, but through Web referrals, happy users and individual donations from around the world.

## **"Those who do not know history...:" VLEs and eHumanities Infrastructure**

Chris Mackie

The international higher education community is engaged at present with the question of how to deliver effectively shared technology services in support of research. Many different approaches can be found, reflecting many different philosophies. We have been here before. Eight to fifteen years ago, higher education was similarly engaged with questions of how to

support eLearning. After a period of innovation and experimentation, two dominant models emerged. Those two models are now both moving into a next generation, being reworked to incorporate new technological possibilities. What lessons can we learn, and then apply to eHumanities support, from the history of eLearning support? In particular, what can the historical trajectories of various eLearning initiatives tell us about optimal ways to support Humanities research, first during the transition to widespread infrastructural support, and then after the infrastructure matures?

## **CLARIN: A Pan-European Research Infrastructure for Language Resources and Technologies**

Martin Wynne

The CLARIN project is a large-scale, pan-European, collaborative effort to support and make more effective the creation, sharing and use of digital language resources and technologies in academic research across the Humanities and Social Sciences. CLARIN aims to lower the technical, legal and administrative barriers to the advanced use of language technologies in the Humanities by building sustainable structures and processes to promote standards, good practice and technical interoperability. A key element in this process is to enable the participation of the Humanities and Social Sciences in the latest national and international developments to build computing and research infrastructure.

## **Scholarsource. A New Paradigm for Digital Humanities**

Paolo D'Iorio

Starting from an analysis of the traditional research infrastructure and its problems, I propose a new paradigm for building digital research infrastructure for the Humanities. This paradigm will 1) take in account the conditions of possibility of Scholarship, 2) rely on "Scholarly Communities on the Web" to lead the way towards digital scholarship, and 3) use some technologies currently developed within the Discovery project ([www.discovery-project.eu](http://www.discovery-project.eu)).

## **Historians, Computer Scientists and the Digital Revolution: Why Collaboration Matters**

Robert N M Watson

Cross-disciplinary collaboration has long been a feature of historical research. Methodologies adopted and adapted from social and cultural anthropology, geography, sociology and literary criticism have transformed history as a field, and allowed scholars to expand their understanding of historical sources and patterns. Yet the digital revolution, while it has been embraced by historians in terms of digital catalogues and online books, has not been subjected to the same level of engagement and methodological debate within history. We argue in this paper for the need for exchange, debate and collaboration among computer scientists, librarians and archivists, and historians to address both the potential and the serious challenges and inequalities of digitization. Historians'

work on the history of the book, on the uses of the archive and on information orders bring a necessary perspective to the construction of new kinds of digital archive. The Digitization of History project at the Centre for History and Economics has hosted a series of interdisciplinary meetings of computer scientists and historians which have highlighted the necessity of collaboration, particularly for the digitization of archival versus printed matter, and for the digitization of non-Roman scripts. We will present a few short case studies to provide examples of the kinds of benefits such a collaboration might bring, and also the kinds of problems it might avoid.

## **From Scriptorium to Collaboratory**

Raphael Lyne

(<http://scriptorium.english.cam.ac.uk/>) aims to digitise a corpus of miscellanies and commonplace books, and to support them with research and teaching resources. It is funded by an AHRC Resource Enhancement grant. Part of its function is to recreate, in digital form, as much of the archival experience as possible: access and conservation are central to the project. It must also anticipate, and work with, the ways in which users and potential collaborators encounter a digital resource. My paper, then, will discuss the ways in which these conceptual and practical borders are crossed, and especially how old and new technologies of reading interact.

## **Mass Digitisation and Visualizing: The Great War Archive**

Alun Edwards

How do you create a community collection of memories and artefacts about the First World War? How do you encourage the public to upload these to an online service? How do you ensure that this collection can become a valuable resource for research and teaching?

This presentation will briefly outline the methods and results of the Great War Archive, a community collection which launched on the 90th anniversary of the Armistice (November 11th 2008). This was a 3-month piece of work which won a 'Highly Commended' for the "Outstanding ICT Initiative" of the year in the 2008 Times Higher Education Awards. It is part of the First World War Poetry Digital Archive from the University of Oxford <http://www.oucs.ox.ac.uk/ww1lit/>. The exciting potential and the pedagogical opportunities of integrating and deploying these digital resources in creative work, teaching and learning, and scholarship are now being realised. The presentation will show the work to investigate and exemplify how existing Web 2.0 tools can be utilised to 'visualise archival data' for these purposes. Techniques being explored include:

- \* geolinking of key historical artefacts in the archive via Google Maps and Google Earth;
- \* podcasting of audio and visual content as an open educational resource for use in teaching via applications such as YouTube!, iTunes and iTunesU;
- \* using collaborative writing tools and feeds such as blogs and wikis;
- \* creating timelines using applications such as Simile;

- \* creating meaningful contexts for the material via social bookmarking sites such as del.icio.us and MyIntute;
- \* releasing the content through federated search engines;
- \* creating informal social networks through Facebook, and Flickr sites;
- \* demonstrating the use of pathways (using tools like the Path Creation Scheme developed at Oxford, LAMS, and VUE).

What will be the perceptions of trust or the validity of these objects once they become disassociated from the main archive? ...this is just one of the questions that this presentation will throw open for discussion!

## **Digital music: Beyond Text?**

Ian Cross

Abstract: The object of study for musicology has traditionally been the musical score and the documents that provide evidence concerning its historical and cultural context. More recently, musicology has begun to turn some of its attention to recorded sound, and some potential benefits of accessing both these types of source materials in virtual form have been recognised. However, music has proven to be rather less susceptible to computational interrogation than have many other foci of research in the humanities. This presentation will briefly sketch the history of digital musicology and will explore some of the problems that surround the virtualisation of music.

## **Cambridge Images Project: A Federated Architecture**

Rod Rivers

Several parts of the university are coordinating efforts over the design, development and implementation of Digital Asset/Image Management Systems. This talk will introduce the Cambridge Images Project and present draft plans for a federated architecture that would facilitate departments and offices in sharing and archiving images while still retaining local systems and controls.

## **Documents, Texts, and Images in a Virtual Research Environment**

Alan Bowman

(<http://bvreh.humanities.ox.ac.uk/VRE-SDM>)

The BVREH programme began with an attempt to discover the what were the priorities of the user communities in ICT application across the range of research represented in the Humanities Division at Oxford. The further aim was develop demonstrators in specific fields of research which would not be driven by technology but would respond to the specific sets of user needs and requirements. The VRE for the Study of Documents and Manuscripts represents a development of such a demonstrator which will show how we can create a virtual laboratory in which scholars working on texts (in any language, culture or

period) can have access to texts of manuscripts and documents, to images of them which can be manipulated and enhanced, to annotation tools, reference materials and other scholarly instrumenta, allowing collaboration in real time with colleagues in remote locations via the Access Grid.

Our current work is also attempting:

- to expand the environment by directly linking textual to archaeological resources, treating documents not just as texts but as artefacts in a physical context
- to improve and customize tools for signal-processing and image-enhancement
- to develop web-based services which articulate and support the cognitive and decision-making processes involved in deciphering difficult texts (see <http://esad.classics.ox.ac.uk>)

## **JISC VRE Programme: The Next Phase**

Frederique van Till

I will be presenting the findings from the projects in the earlier phases of the Virtual research environments programme as well as the plans for the upcoming third phase, that will start off in March 2009.

The JISC VRE programme has been running since Spring 2004 and has now reached a phase of consolidation and redeployment of solutions within the wider academic community. The approaches in which this will be done complement the developments of the Bamboo project and Digital Humanities.

## **ADONIS-tge**

Yannick Maignien

ADONIS-tge is a Major Facility (*Très Grand Équipement*) launched by the CNRS (French National Centre for Scientific Research) open to other partners. It aims to promote integrated access to digital data and documents in the Humanities and Social Sciences.

What does 'Integrated access' or 'Unified access' to the extreme diversity of disciplines, digital objects, various institutions and many uses mean ?

We try to focus on interaction between

- Access to data and search engines
- Managing metadatas
- Archiving
  - Cooperation on data processing
  - Publishing
  - Software open repositories solutions

The key issue is the notion of Meta-portal, as a convergence of these constraints of access, preservation, publications and uses ...In other terms : Building a E-Science or CyberInfrastructure, for SHS purpose.

We think that we know very little about the real interactions between this "cycle" from production to communication inside social sciences.

Our concern is to be open to new relationships and new programs in contemporary social sciences and humanities : to the extreme complexity of interdisciplinary projects, even with nature's sciences (health, warm climate, neuro-sciences, immigration, etc...)

The notion of Meta-portal means Interconnecting data, metadata, but also players and services ...And these two last items are the biggest issues!

I would particularly stress the difficulties involved in the "social networking" part of the Adonis project rather than the technical ones (even if assessment of Fedora Commons Solutions for such an integrated project represents an enormous challenge).

How social and intellectual institutions would evolve when faced with this technical and computational change seems to be a key factor for this kind of project. This is also why we need straight cooperation with other similar projects around the world.

## **Bamboo: Defining Infrastructure, Creating a Community**

Chad Kainz

Funded by the Andrew W. Mellon foundation, Project Bamboo is an 18-month arts and humanities cyberinfrastructure planning project led by the University of Chicago and University of California, Berkeley, that aims to explore how shared technology services can enhance research, learning, and scholarship. The presentation will provide a brief planning project update, and share ideas of how Bamboo may contribute to cross-institutional and cross-disciplinary academic communities, promote scholarly collaboration, and foster the development of sustainable projects and activities.